Contributions to Visual Autonomous Driving

A Review

Ernst D. Dickmanns

Part I: Basic Approach to Real-time Computer Vision with Spatiotemporal Models (1977 ÷ 1989)

Abstract (Part I)**:** After a very brief survey on the development of vehicles and microelectronic devices like video cameras and digital microprocessors the first steps in developing vehicles with the sense of vision are reviewed. The paper then concentrates on digital onboard processing for road vehicles and the breakthrough brought about by the 4-D approach introducing spatio-temporal models for image sequence processing. A first cooperation with Daimler-Benz led to the definition of a specific sub-project in 'Pro-Art' of the EUREKA-project PROMETHEUS to be discussed in Part II; this has shifted weights in international developments of machine vision.

I.1 Introduction

Ground vehicles just had wheels at the tips of one or two axles for about five to seven thousand years; there is some uncertainty arising from first pictures found on a vase and from the remaining of real wheels excavated. Animals as source of energy for locomotion have also been in use for thousands of years; by their sense of vision and their memory of the environment around home they even had a limited capability of autonomous driving. Especially horses found their way home in road networks when the human navigator had fallen asleep or was distracted otherwise. – Engines replacing animals are around for a few hundred years only. Light-weight engines for wheeled ground vehicles are in use for just about 130 years. Humans were willing to trade the limited autonomy of animals for the need of steady attention while driving at higher speeds also over longer distances. – Video cameras as technical equivalents of the biological eye are available since less than a century; initially, these cameras have been very bulky and needed much electric power. Only the introduction of transistors that led to microprocessors (since the 1970s) generated the preconditions for small mobile systems needing low energy levels only. 16-bit microprocessors became available on the market in the early 1980s – together with the 'Personal Computer' (PC). This allowed, for the first time, to think realistically about providing the sense of vision for road vehicles by flexible digital data processing on board.

This article is intended to give a detailed survey on the efforts at the University of the Federal Armed Forces in Munich (UniBwM) where a unique approach has been developed that dominated the first decades of autonomous driving with regard to the ratio of performance / cost. A few cross references to activities world-wide will be given every now and then to allow a wider perspective. Surveys on the history of autonomous driving may be found in [Tsugawa 1994 (for Japan); Bertozzi et al. 2000; Dickmanns 2002; Weber 2015].

Some experiments in the direction of visual autonomous driving had begun in the 1960s [Nilsson 1969; Gennery 1977; Moravec 1979]; all vehicles in these studies worked intermittently with the digital computers needed in a remote laboratory, off-board. Tsugawa in the 1970's was the first to investigate visual guidance of a road vehicle with signal processing on board (partly analog, using hard-wired logic and a simple guidance scheme); he confined the study to tracking lateral guide rails by a vertically arranged pair of stereo video cameras on a real car. The speed of the vehicle was about 10 km/h over a short distance [Tsugawa et al.1979].

I.2 Early visual autonomous driving with digital processors onboard

This development started in Germany and in the USA independently and separately without the parties involved knowing of each other; it was only in 1984 at a "NATO Advanced Research Workshop on Vision and Image Understanding" in Erice, Sicily, that the groups learned to know of each other superficially. Both had thought to be the only ones working in this direction. DARPA had held its project ALV (see below) secret, and the report [Dickmanns 1980] was internal only. Relevant publications first hit each other at the SPIE-Conference on 'Mobile Robots' in Cambridge, MA in the fall of 1986 [Davis et al. 1986; Dickmanns and Zapp 1986; Kuan et al. 1986; Thorpe and Kanade 1986]. While the approach in Germany was a singular activity initiated by Dickmanns, the American one was part of a large national effort in the USA for developing the next-generation real-time computers ("Strategic Computing", [Roland and Shiman 2002]) with massively parallel processors (MPP). More than a dozen proposals have been funded by the **D**efense **A**dvanced **R**esearch **P**roject **A**gency (DARPA) initially; three application areas were to be studied: 1. Perceiving the environment of a carrier vessel including moving objects both under water, at the sea surface, and in the air. 2. An 'Autonomous Autopilot' for a fighter aircraft, and 3. an **'Autonomous Land Vehicle' (ALV)**, of special interest here. The University of Maryland in College Park (Prof.s I. Rosenfeld, L. Davis), Carnegie Mellon University in Pittsburgh (Prof. T. Kanade, C. Thorpe), and the University of Massachusetts in Amherst (Prof.s A.R. Hansen, E. Riseman) were main contenders among several others.

I.2.1 The general approach chosen by UniBwM

As acting head of the research center Oberpfaffenhofen of DFVLR (the German Research and Test Facilities for Aero-Space Technology, now DLR) Ernst Dickmanns in 1974/75 had become acquainted with the state of the art in digital processing of snapshots taken by airplanes and satellites, on the one side, and with the general German policy of federal funding of projects, on the other side: Industry projects had to have a short look-ahead range of a few years only; projects of the institutions for applied research (like DFVLR) should have a look-ahead range of up to about ten years, and institutions for basic research like Max-Planck-Institutes and universities should concentrate on longer range developments of technology beyond a decade away from application.

When he was called to the newly founded university of the Federal Armed Forces in Munich (UniBwM, Neubiberg) in 1975 for the rest of his professional career, he combined this knowledge with predictions of future developments in digital micro-processing. The experience in this field showed that an order of magnitude in computing performance should become available every four to five years in the same framework of microprocessors. This span happened to be the time usually needed for achieving a Ph.D.-degree at a German university. So, within two to three decades (about five Ph.D.-generations) a factor of 10^6 (one million!) should became available driven just by market forces. Looking at the processing times actually needed for image interpretation led to the conclusion, that real-time image sequence evaluation at video rate by microprocessors should become possible within the next decades. With the background of control engineering for dynamical systems he decided to pick '**real-time vision**' for guidance of any type of vehicles as research field. Since motion control in all six degrees of freedom - as needed in aero-space applications - was too complex for the beginning, he developed a stepwise plan to approach this final goal [Dickmanns 1980].

The **introductory step** was a highly dynamical task with **one degree of freedom** only: **Balancing a rod on an electro-cart** with acceleration / deceleration in one translational degree of freedom (*d.o.f.*) as the only control variable. Such a vehicle was available on the market for experimental studies in control engineering for students in higher semesters. The vehicle was able to accelerate at a rate of ~ 8 m/s² (0.8 of Earth gravity "g"). Combining it with a rod supported at its lower end by a joint with one rotational *d.o.f.* orthonormal to the direction of motion of the cart yields a very dynamic mechanical system. The length of the rod determines the eigen-frequency of the combined system; the shorter the rod, the higher is the eigen-frequency. Rod lengths between 0.5 and 2 m have been used [Meissner 1982; Meissner and Dickmanns1983]. The demonstration of this system at an external conference in 1982 probably was the first really dynamic scene controlled by real-time computer vision realized on half a dozen 8-bit micro-processors. Details including a video clip may be found in (www.dyna-vision.de section **3.1.1**).

The **second step** involved **three *d.o.f.*** in slow **planar motion**: Contrary to the highly dynamic pole balancing task (especially with small lengths of the pole), the satellite model plant is characterized by sluggish motion controlled by bang-bang digital input (valves closed or open, 1983 – 1987). Two pairs of thrusters at opposite sides of an air-cushion vehicle allow almost frictionless control of two translational and one rotational *d.o.f.*; the actual rotation angle determines the thrust direction. In each *d.o.f.* **two simple integration steps** connect the control input to the state variables of the object; these dynamical models constitute core knowledge about the motion process observed. Nonlinear perspective mapping of corner features on the objects observed constitutes the first part of the measurement process yielding image features; measuring their position in the image constitutes the second part of it. Note that the states measured are not the positions of the features in the image but the actual physical states of the real vehicle. Thus, in recursive estimation using dynamical models, velocity components are reconstructed as a byproduct by integration of prediction errors of feature positions; differencing of two consecutive pose states for velocity estimation, as done usually and leading to noise amplification, is avoided. Details including a video clip may be found in (www.dyna-vision.de section **3.1.2**).

Both results have been achieved with about half a dozen conventional 8-bit microprocessors in a custom-designed processing system "**BVV1**". Since at that time it was impossible to read and evaluate full (even small 240 x 240 pixel) video images in real-time, a 'window-approach' had been designed by Dickmanns and realized by Graefe [Dickmanns and Graefe 1988]. Up to six arbitrarily locatable sub-windows of size 32 x 32 pixels could be captured from a standard 'video-**field**' (either all even or all odd lines forming the video-**half-images**); this reduced the data volume by a factor of two and provided an additional half-frame time for data processing. Storing pixels from previous images, which made other wide spread approaches to image sequence processing so demanding, was completely avoided by adapting the well-known method in control engineering 'prediction-error-feedback' (Kalman filter [Kalman 1960], Luenberger observer [Luenberger 1964]) to image sequence processing. Perspective projection was considered in this approach to be a nonlinear measurement process reducing the 3-D spatial world to a 2-D image sequence; so image data were not considered to be 'measured states' of a dynamical system (as was usual in the Computer Science / Artificial Intelligence approaches at that time). In these approaches, time-differencing (of course, with noise amplification) of information obtained by *inverse* perspective projection from consecutive images, was the step used for obtaining data on the real-world-system. On the contrary, in the so called "4-D"-approach of Dickmanns (3-D space plus time) an image was just an intermediate frame for data storing; the **actual true physical states of the dynamical system** observed have been used as primary variables. The nonlinear transformations of perspective projection have been dealt with through linearization of the measurement equations. In order to keep linearization-errors small, a high image evaluation rate was mandatory. Therefore, quite contrary to the approaches used elsewhere in computer vision, in the UniBwM-group the side constraint was introduced that image evaluation frequency should not drop below ~ 10 Hz (~100 milliseconds (ms) per image) [Dickmanns 2007 (Chap. 7 to 9); from here on abbreviated as 'Di.07 Chap.7-9']. Almost all other approaches worldwide worked with many seconds or even minutes for a single frame, hoping to achieve real-time performance by increased computing power in the future. These considerations have led to the sequence of steps partially mentioned above:

1.) **One**-*d.o.f.* pole balancing using **edge features**;

2.) **Three**-*d.o.f.* slow 'satellite docking' using **corner features** and a sluggish air-cushion vehicle with bang-bang control;

3.) **Three**-*d.o.f.* road vehicle guidance with three relatively simple continuous control functions in an environment perceived through **extremely variable operators** for **edge features**;

4.) **Six**-*d.o.f.* landing approach of: 4a) aircraft and 4b) helicopter with several control functions that interact in a complex way.

Visual guidance of road vehicles has been prepared by using the simulation facilities specially designed for the sequence of tasks mentioned under points 3 and 4 above: In the field of general aero-space engineering, simulation loops for training human pilots on the ground using visual feedback had been developed over the last decades. Taking advantage of this technology available, the funding for the newly founded university was used to design such a **simulation loop for machine vision systems** capable of **driving cars** and **flying aircraft** (Fig. I.1). Both, rather dynamic simulation of real motion in three rotational *d.o.f.* for a real sensor package with inertial and visual components, and a cylindrical projection screen for visual display including the components for computational graphics has been realized (see

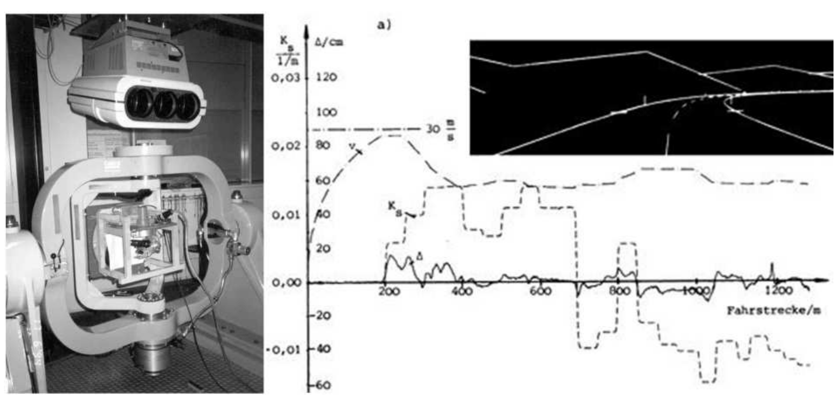

Figure I.1: Early Hardware-In-the-Loop (HIL) simulation of ground vehicle guidance by computer vision (starting 1979) www.dyna-vision.de section **3.3.4**). The then newest (moderately priced) calligraphic display system “Picture System 2” used hardware especially designed for quaternions and *generalized coordinates* with four by four matrix operations for the first time [Roberts 1965]. The advantages convinced the UniBwM-group to also adopt this framework as general background for image sequence analysis. H. G. Meissner was the first to apply this simulation loop to the task of road vehicle guidance using calligraphic (vector) pictures in closed-loop form, both in display and in the evaluation of corresponding video sequences [Meissner 1982] (see Fig. I.1). The first international publication on road vehicle guidance with results from this simulation loop was [Dickmanns and Zapp 1985]

I.2.1.1 The 4-D approach to autonomous visual guidance of vehicles (1977 ÷ 1986)

While known visual guidance schemes for ground vehicles dealt with small, slowly moving devices that either stood still and observed the environment or moved blindly in laboratory- (Stanford cart [Nilsson 1969; Moravec 1979, 1983] or planetary environments [Gennery 1977], Dickmanns had the idea of choosing an application area that was visually simple by definition, yet of real practical value right from the beginning. Using the Autobahn with well-defined construction parameters and a reduced set of objects allowed on it (no animals and no humans on foot or on bicycle, no crossroads and no traffic lights) in daily commutes, this seemed to be an ideally suited environment for demonstrating the practical advantages of machine vision. The required relatively high speed (at least 40 km/h in Germany) was considered to pose no difficulty since at 10 Hz image evaluation rate and a speed of 36 km/h (10 m/s) {108 km/h (30 m/s)} the vehicle would move only 1 m {3 m} from frame to frame. With a look-ahead range of ~ 30 m, one specific slice of road between 5 and 30 m may thus be tracked in the image sequence 25 {8} times. This is ideally suited for understanding scenes in which curves (lane markings, road boundaries) follow generic laws: In Germany, clothoids with piecewise linearly changing ‘curvature over arc length’ are used for high-speed roads.

The positive results of Meissner and Zapp in simulation led to the acquisition of a 5-ton van in 1984 and it’s equipping with the necessary modifications by a small company near the university, also from initial funding of the institute. There was no involvement of Daimler-Benz AG (DBAG) whatsoever. The “**V**ersuchs-fahrzeug für **a**utonome **Mo**bilität und **R**echner**s**ehen” (experimental vehicle for autonomous mobility and computer vision) – dubbed **VaMoRs** – became available for first tests in 1985 (see Fig. I.2); dynamical models for describing its mobility parameters in a knowledge base have been determined through theses-work of students till mid-1986. As far as is known, at that time this was the only vehicle on the globe with dynamical models used for longitudinal and lateral

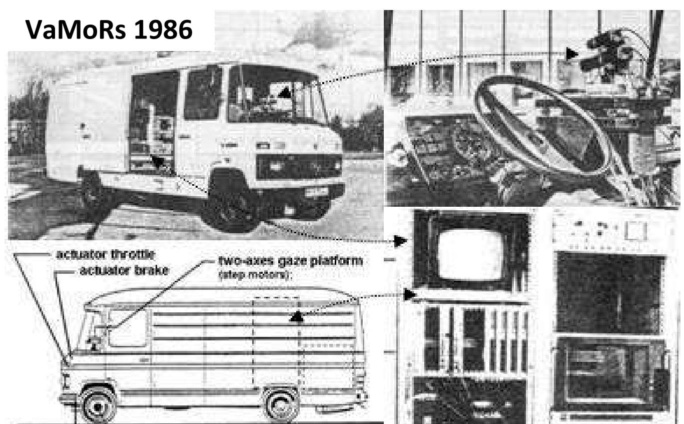

Figure I.2: Original version of VaMoRs 1986; a 5-ton van with additional 220 V electric power generator, a standard 19’’ industry rack for electronic devices, and active gaze control for a bifocal camera set (top right).

control in a knowledge base for visual autonomous guidance.

I.2.1.2 First publicly funded projects (1982 ÷ 1991)

In the framework of a 'Funding Program for Information Technology' of the 'Federal Ministry for Research and Technology (BMFT), UniBwM received funding from 1982 to 1984 for 'Using models to improve the interpretation of dynamical scenes' in the field of visual autonomous road vehicle guidance. This was followed by two contracts from the German Science Foundation (DFG) for 1.) "Perception and control of motion in a technical environment by processing of image sequences" from 1984 to 1987; satellite docking and road vehicle guidance were the application areas. 2.) "4-D-Scene recognition with integral spatiotemporal models" from 1987 till 1991; here, beside guidance of ground vehicles, autonomous visual landing approaches have been investigated.

One of the essential results of this phase was to switch to models from differential geometry for describing the shape of roads and lanes. Since only local curvature of the road and lateral position on it is of importance for just driving properly (no larger-scale navigation goals), the absolute geometric coordinates may be neglected. This reduces the number of unknown road parameters to just two in a planar environment (local curvature C_0 and curvature change with arc length C_1, see Fig. I.3). As will be seen later, this also is the base for simple superposition of vertical curvature in order to describe road geometry with sufficient accuracy in hilly terrain.

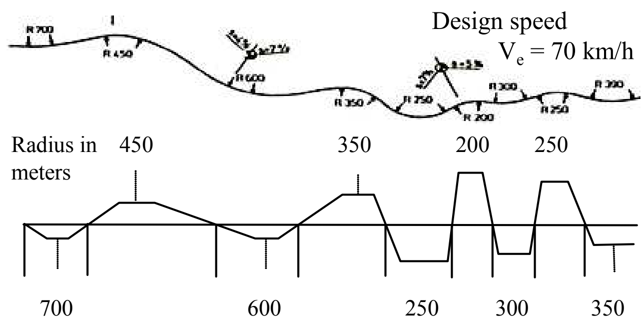

Figure I.3. Road (upper curve in a bird's-eye view) as a sequence of clothoid elements with continuity in position, heading, and curvature; the curvature change rate C_1 is a sequence of step functions. The curvature over the arc length is a polygon (lower curve).

Another long lasting result was the general architecture for the very efficient 'window-

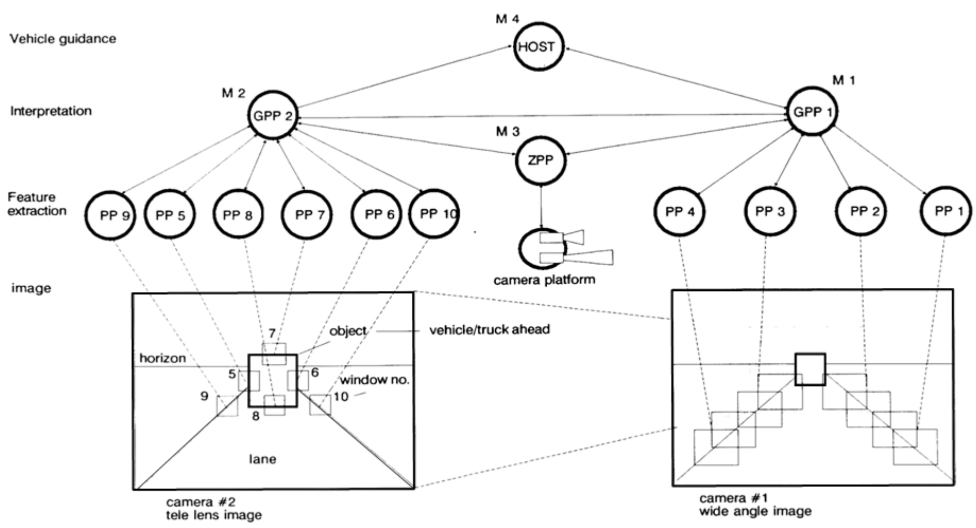

Figure I.4: First-generation system architecture for dynamic vision in road scenes of the 1980s; the parallel processors for edge feature extraction (PPi) were 16-bit Intel 8086 microprocessors capable of extracting just a few edge features per observation cycle of 80 ms.

approach' shown in Fig. I.4 (see [Dickmanns, Christians 1989/91]). The functionality of the custom-designed microprocessor system BVV2 with about a dozen 16-bit units Intel 80X86 already displays the three levels in visual dynamic scene understanding: The lowest one is oriented towards image features with temporal continuity; the second one is grouped around real objects in the scene observed. Vehicle guidance on the third level is just a seed item to grow into situation assessment in further developed systems of later years. Feature extraction was based on very flexible operators for finding arbitrarily oblique edge features with adjacent average gray values from video-fields (half-images) [Kuhnert 1988; Mysliwetz 1990; Di.07, Section 5.2]. Quadratic subfields of 32 x 32 pixels (lower part of Fig. I.4) had been dynamically grabbed while flying by; only these subfields have been stored and evaluated. No storage of entire images or fields was necessary. The result of all previous evaluations was captured in the estimated actual state variables of the models used. In the next image of the sequence edges with neighboring orientations were searched for in the same region.

I.2.2 First cooperation with Daimler-Benz (1986 ÷ 1989)

After first publications on visual pole balancing and satellite docking, where road vehicle guidance had been mentioned as a running activity, management personnel from Daimler-Benz AG (DBAG) approached UniBwM whether they would be willing to do a joint project in road vehicle guidance. In 1986 BMFT had announced funding for joint projects between industry and universities. This resulted in the project "Autonom Mobile Systeme" (AMS). In order to get the project approved by the director for 'Research' of DBAG, UniBwM had to do a demonstration in December 1986 in the DBAG-skidpan in Stuttgart (www.dyna-vision.de section **3.1.3.6**). **VaMoRs** accelerated along a straight line up to 10 m/s (36 km/h). Then a spiral curve with increasing curvature started that ended almost tangentially to a circular border line in the skidpan formed by surface material of different gray values (concrete and cobble stones). **VaMoRs** then had to adjust its speed such that lateral acceleration was 1 m/s²; this led to a reduced speed of about 7 m/s at which it then circled in the skidpan. Both longitudinal and lateral guidance have been done by computer vision right from beginning [Di.07, Section 7.3.3].

Goal of the sub-project in AMS ("Improvement of traffic safety by automatic recognition of obstacles and vehicle guidance") that ran from 1986 till 1989 was to equip a 'Spurbus' (a large bus, see Fig. I.5) available at DBAG from a previous project with our vision system and to:

1.) Demonstrate visual guidance along the entire test track at the DBAG-test-site in Rastatt, and
2.) to show the capability of stopping in front of a stationary object (garbage can of size ~ 0.5 x 1 m) from an initial speed of 40 km/h.

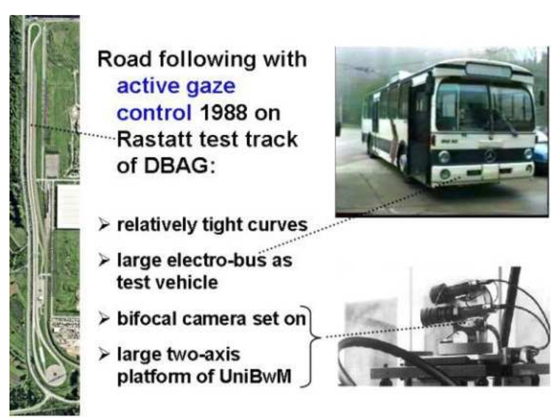

Figure I.5: Final demonstration 1988 of project AMS at the Daimler-Benz test site in Rastatt.

With the 4-D architecture both goals of the project were achieved as scheduled and demonstrated in 1988 (www.dyna-vision.de section **3.1.3.2**).

When the director of the 'Autobahn-Direktion Süd' in Munich heard of the project he informed UniBwM about the fact that there was a new piece of Autobahn of more than 20 km in length near Dingolfing not yet turned over to the public because the finalization of a bridge was lagging in time. He offered this route with finished surface but partially missing lane markings to be used as test track for **VaMoRs**. UniBwM gladly accepted this offer, and in the spring and summer of 1987 A. Zapp and his coworkers were able to increase the maximum speed of autonomous visual driving to the maximum speed of the vehicle of 96 km/h (60 mph), limited only by engine power [Dickmanns and Zapp 1987; Zapp 1988] (www.dyna-vision.de section **3.1.3.1**). This was about an order of magnitude better performance than demonstrated by any other vehicle around the globe guided by machine vision. {Newspaper reports: **Neue Zürcher Zeitung** dated 21.4.1987; **Weltbild** dated 30.12.1987. Journals: **Hobby** 3/88, 1988; Hochschulkurier UniBw Munich 14/88, 1988; **National Geographic** June 1989; TV-broadcast **BBC** in the series ‚**Tomorrow's World**' 1988: Self-Drive-Van}.

These results and those of [Wuensche 1986; 1987] led to the first publication on the so-called "4-D approach" to real-time computer vision exploiting spatiotemporal (dynamical) models [Dickmanns 1987b; Di.07, Chap. 6]. This new method – an extension of recursive estimation, well known in control engineering, to perspective projection and image sequence evaluation – brought about a quantum jump in performance with off-the-shelf microprocessors; however, they had to be arranged and triggered in a custom-designed approach realized in the BVV2 [Dickmanns and Graefe 1988]. The performance achievable in different domains became known to a wider audience by a keynote talk at IJCAI 1989 [Dickmanns 1989] and the videos shown.

I.2.3 Definition of the PROMETHEUS-project (1986 / 87)

In the mid-1980s DBAG had started preparing a large European research effort for the second century in the development of road vehicles (referred to the first automobiles by K. Benz and G. Daimler in 1886). In this EUREKA-project "**PRO**gra**M**me for a **E**uropean **T**raffic of **H**ighest **E**fficiency and **U**nprecedented **S**afety" (PROMETHEUS), among many other things DBAG had proposed to develop autonomous lateral guidance of cars using electromagnetic fields generated by cables buried in the center of the lanes. Noting that this would introduce for these cars the disadvantages of trains rigidly guided laterally by rails, Dickmanns proposed to substitute this plan by a big European effort in machine vision which would allow, beside lateral guidance, simultaneous detection and avoidance of obstacles as well as exploiting also for the autonomous system (at least in the long run) much of the infrastructure developed for communication with human drivers. Lane changes and handling of intersections or road forks would also be possible without additional costs in infrastructure as would be needed with buried cables and inductive electric fields [Dickmanns 1986].

The higher management levels in the automotive industry remained skeptical with respect to autonomous driving for several years, but agreed to study computer vision for driver assistance functions [Braess und Reichart 1995]. On the contrary, the goal of the UniBwM group followed by several other European participants in the sub-project 'Pro-Art' (for 'Artificial Intelligence') was to develop the capability of autonomous driving right from the beginning [Dickmanns et al. 1987]; the argument was: How can one optimally assist somebody without knowing how the matter of assistance is done properly. In addition, aspects of autonomous driving were being developed in a European cooperation by funding from the German Ministry of Defense anyway [Dickmanns 1987c].

It took some time to get through with this proposal in the automotive industry, but the results with **VaMoRs** in 1987 and the DARPA-effort in the USA finally convinced all European participants. More than a dozen European car manufacturers and about five times that number of European universities embarked on this effort for the next seven years.

I.3 A comparative look at the development in the USA

Triggered by the Japanese effort towards fifth-generation computers, in the early 1980s the US-DARPA launched a development program for preparing the next-generation of real-time computers ("Strategic Computing") [Klass 1985]. Computer vision was to be an essential component; therefore, derived from biological vision systems with massively parallel signal processing starting right after signal acquisition immediately in the eye, an essential component of the system architectures investigated was massively parallel processing (MPP) with up to millions of relatively simple units [Weems et al. 1990; Hillis 1992]. More than a dozen proposals for architectures have been funded by DARPA initially; three application areas were to be studied (see introduction to section I.2 above).

The ALV-program ran from 1982 till 1989 with the industrial company Martin-Marietta in Denver, Colorado responsible for the official test vehicle 'ALV', an 8-ton van with four rigid axles, each with twin wheels at each tip). Heading was controlled by different wheel speeds on each side (like a tank with tracks). This was clearly intended for cross-country driving at slow speeds. The development of various software systems was done at several universities as mentioned above; first test drives have been done on roads in 1986. There were no safety personnel on board during test rides; they had to follow the vehicle on foot and were expected to shut down the autonomous system by an outside button in case something unexpected happened. The internal room of the vehicle was full of power generators, sensors, and computer hardware as well as space for programmers when parked. Since the computer systems to be specially developed for vision were not yet available in the beginning, modified standard processors were used initially for testing various approaches; they all needed many seconds to minutes for evaluating one full video image. Inverse perspective projection of consecutive images has been used for arriving at estimates for the environmental parameters and relative state of the vehicle.

The Robotics Institute at Carnegie Mellon University (CMU), where H. Moravec had been working with small laboratory vehicles in continuation of the early Stanford efforts, decided for a standard van of their own, equipped with a special reduction gear for slow driving and with a set of powerful parallel microprocessors in the rear space [Kanade et al. 1986]. There was the standard seat for a safety driver behind the steering wheel and room for programmers also while driving very slowly. Initial tests have been done on roads with trees near to the driveway that were reported to pose special difficulties with the shadows they generated.

Starting in 1985, results in autonomous driving with all vehicles mentioned above were presented at yearly international conferences like IEEE "Robotics and Automation" (ICRA) and SPIE "Mobile Robots". Since 1992 the yearly specific International Symposium on "Intelligent Vehicles", the location of which rotated yearly between the USA, Europe, and Asia (in Asia initially Japan only), provides a good source for judging progress. A survey on the state of the art in 1987 has been given for the Prometheus-community in [Dickmanns 1987a]. Relatively long image evaluation times and slow driving speeds predominated well into the 1990s.

For obstacle detection, additional active sensors based on radar or 'laser range finders' (LRF or lidar: laser induced detection and ranging) predominated; much effort went into developing corresponding active sensors around the globe. Monocular stereo-vision for range estimation to other objects was not considered to be viable. It was the 4-D approach that allowed monocular vision-only depth estimation by taking motion over time and odometer signals into account [Dickmanns and Christians 1989]. This has not been adopted by American colleagues until the final Prometheus-results were published (see Part II). One specific revolving laser range finder with

360° field of view was the outcome of this kind of thinking; it was to shape an appreciable part of US-developments in autonomous driving that will be reviewed towards the end of Part III.

I.4 Conclusions (of Part I)

The 4-D approach using spatiotemporal models and recursive estimation with feedback of prediction errors was proven to be very efficient. Around the 1990s, this approach was still in strong competition with neural network approaches and those relying on massively parallel (simple) processors under study at many institutions around the world. However, with the progress in the development of general-purpose microprocessors and of communication networks, UniBwM saw no reason to change, except for the newly developed microprocessors with four direct communication links to its neighbors: the ‘transputers’. This has led to the development of the second-generation vision system of UniBwM discussed in Part II.

References (Part I)

Braess, H. H. u. Reichart, G.: Prometheus: Vision des «intelligenten Automobils» auf «intelligenter Strasse»? Versuch einer kritischen Würdigung. Teil 1. ATZ Automobiltechnische Zeitschrift 97, Nr. 4, S. 200-205; Teil 2. ATZ Automobiltechnische Zeitschrift (1995) Nr. 6, S. 330–343

Bertozzi M., Broggi A., Fascioli A. 2000: Vision-based intelligent vehicles: State of the art and perspectives. Robotics and Autonomous Systems 32: 1–16

Davis L., Kushner T.R., Le Moigne J.J., Waxman A.M. 1986: Road Boundary Detection for Autonomous Vehicle Navigation. Optical Engineering, 25(3): 409–414

Dickmanns E.D. 1980: Untersuchung und mögliche Arbeitsschritte zum Thema „Künstliche Intelligenz: Rechnersehen und –steuerung dynamischer Systeme“. HSBWM / LRT / WE 13 / IB 80-1 (18 pages)

Dickmanns E.D., Zapp A. 1985: Guiding Land Vehicles Along Roadways by Computer Vision. Proc. Congres Automatique 1985, AFCET, Toulouse, pp 233-244

Dickmanns E.D., Zapp A. 1986: A Curvature-based Scheme for Improving Road Vehicle Guidance by Computer Vision. In: 'Mobile Robots', SPIE Proc. Vol. 727, Cambridge, MA, pp 161-168

Dickmanns E.D. 1986: Computer Vision in Road Vehicles – Chances and Problems. ICTS-Symposium on "Human Factors Technology for Next-Generation Transportation Vehicles", Amalfi, Italy

“ 1987a: State of the art review for ‘PRO-ART, 11 170 Integrated approaches’. UniBw M/LRT/WE 13/IB/87-2 (limited distribution) .pdf

“ 1987b: 4-D-Dynamic Scene Analysis with Integral Spatio-Temporal Models. 4th Int. Symp. on Robotics Research, Santa Cruz. In: Bolles RC; Roth B 1988. Robotics Research, MIT Press, Cambridge, pp 311-318.

“ 1987c: Object Recognition and Real-Time Relative State Estimation Under Egomotion. NATO Advanced Study Institute, Maratea, Italy (Aug./Sept) In: A.K. Jain (ed): Real-Time Object Measurement and Classification. Springer-Verlag, Berlin, 1988, pp 41-56

Dickmanns E.D., Zapp A. 1987: Autonomous High Speed Road Vehicle Guidance by Computer Vision. IFAC World Congress Munich, Preprint

Dickmanns E.D.; Graefe V.; Niegel W. 1987: Abschlussbericht Definitionsphase PROMETHEUS Pro-Art, pp. 1-18

Dickmanns E.D.; Graefe V. 1988: a) Dynamic monocular machine vision. Machine Vision and Applications, Springer International, Vol. 1, pp 223-240. b) Applications of dynamic monocular machine vision. (ibid), pp 241-261

Dickmanns E.D. 1989. Detroit, Aug. 23. Invited keynote talk, **IJCAI** on **‚The 4-D approach to real-time vision’** (11:10 – 12:40 Uhr); including several videos with experimental results.

Dickmanns E.D., Christians T. 1989: Relative 3-D-state Estimation for Autonomous Visual Guidance of Road Vehicles. In T. Kanade et al (eds): 'Intelligent Autonomous Systems 2', Amsterdam, Dec. Vol. 2, pp 683-693; also appeared in: Robotics and Autonomous Systems 7 (1991), Elsevier Science Publ., pp 113-123

Dickmanns E.D 2002. Vision for ground vehicles: history and prospects. Int. J. of Vehicle Autonomous Systems (IJVAS), Vol.1, No.1, pp. 1 – 44.

[Di.07] ≡ Dickmanns E.D 2007. Dynamic Vision for Perception and Control of Motion. Springer-Verlag, (474 pp.)

Gennery DB 1977: A Stereo Vision System for an Autonomous Vehicle. Proc. 5th IJCAI, Cambridge, MA, USA.

Hillis W.D. 1992 (6th printing): The Connection Machine. MIT Press, Cambridge, MA

Kanade T.; Thorpe C.E.; Whittaker W. 1986: Autonomous Land Vehicle Project at CMU. Proc. 1986 ACM Comp-Conf., Cincinnatti, Febr. 1986

Kalman R.D. 1960: A new approach to linear filtering and prediction problems. Trans. ASME, Series D, Journal of Basic Engineering: 35–45

Klass P.J. 1985: DARPA Envisions New Generation of Machine Intelligence. Aviation Week & Space Technology, April: 47–54

Kuan D., Phipps G., Hsueh A.C. (1986): A real time road following vision system for autonomous vehicles. Proc. SPIE Mobile Robots Conf., 727, Cambridge MA: 152–160

Kuhnert K.D. 1988: Zur Echtzeit-Bildfolgenanalyse mit Vorwissen. Diss. UniBw Munich, LRT

Luenberger D.G. 1964: Observing the state of a linear system. IEEE Trans. on Military Electronics 8: 74–80

Meissner HG 1982: Steuerung dynamischer Systeme aufgrund bildhafter Informationen. Diss., UniBw Munich, LRT

Meissner HG; Dickmanns ED 1983: Control of an Unstable Plant by Computer Vision. In T.S. Huang (ed): Image Sequence Processing and Dynamic Scene Analysis. Springer-Verlag, Berlin, pp 532-548

Moravec H. 1979: Visual Mapping by a Robot Rover. Proc. IJCAI 1979: 598–600

Moravec H. 1983: The Stanford Cart and the CME Rover. PIEEE(71), 7: 872–884

Mysliwetz B. 1990: Parallelrechnerbasierte Bildfolgeninterpretation zur autonomen Fahrzeugführung. Diss. UniBw Munich, LRT

Nilsson N.J. 1969: A Mobile Automaton: An Application of Artificial Intelligence. Proceedings International Joint Conference on Artificial Intelligence (IJCAI): 509–521

Roberts L.G. 1965: Homogeneous matrix representation and manipulation of n-dimensional constructs. MS-1405, Lincoln Laboratory, MIT

Roland A., Shiman P. 2002: Strategic Computing. *DARPA and the quest for machine Intelligence, 1983 – 1993*. MIT Press, Cambridge, MA; London, England (pp 427)

Thorpe C., Kanade T. 1986: Vision and Navigation for the CMU Navlab. In: SPIE Conf. 727 on 'Mobile Robots', Cambridge, MA

Tsugawa S., Yatabe T., Hirose T., Matsumoto S. 1979: An Automobile with Artificial Intelligence. Proc. 6th IJCAI, Tokyo: 893-895

Tsugawa S. 1994: Vision-Based Vehicles in Japan: Machine Vision Svstems and Driving Control Svstems. IEEE Transactions on Industrial Electronics, Vol. 41, No. 4, August

Weber M. 2015: Where to? A History of Autonomous Vehicles. Computer History Museum. http://www.computerhistory.org/atchm/where-to-a-history-of-autonomous-vehicles

Weems CC, Rana D, Hanson AR, Riseman EM, Shu DB, Nash JG 1990: An Overview of Architecture Research for Image Understanding at the University of Massachusetts. Proc. ICPR 1990, Vol-II, pp. 379-384.

Wünsche H.-J. 1986: Detection and Control of Mobile Robot Motion by Real-Time Computer Vision. In: Marquino N. (ed): Advances in Intelligent Robotics Systems. Proc. SPIE, 727: 100–109

Wünsche H.-J. 1987: Bewegungssteuerung durch Rechnersehen. Diss. UniBw Munich, LRT. Also as Fachberichte Messen, Steuern, Regeln Bd. 10, Springer-Verlag, Berlin, 1988

Zapp A. 1988: Automatische Straßenfahrzeugführung durch Rechnersehen, Diss., UniBw Munich, LRT

Zimdahl W., Rackow I., Wilm T. 1986: OPTOPILOT – ein Forschungsansatz zur Spurerkennung und Spurführung bei Straßenfahrzeugen. VDI Berichte Nr. 162, pp49-60

www.dyna-vision.de : Website with many details about the development of real-time machine vision systems at UniBw Munich, including dozens of video-clips in various applications.